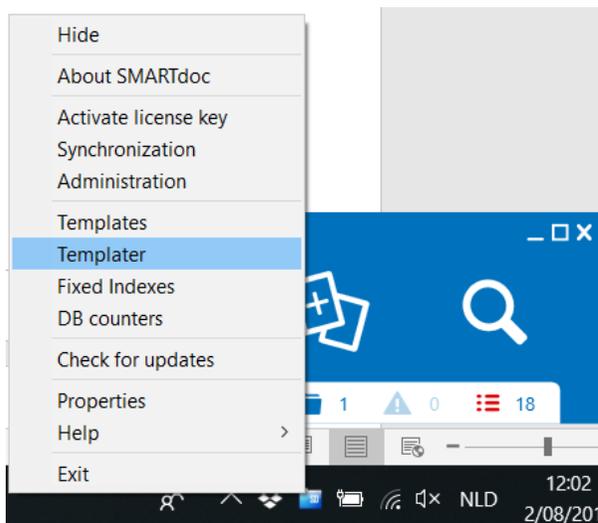


SMARTdoc Templater: Recognition Templates

THE IDEA

The Templater is the tool used to create all kind of SMARTdoc templates used to recognize documents, to create workflows and to be sure our data are well protected by creating the right security templates.

To start the Templater, right click the SMARTdoc icon in the taskbar and choose 'Templater'.



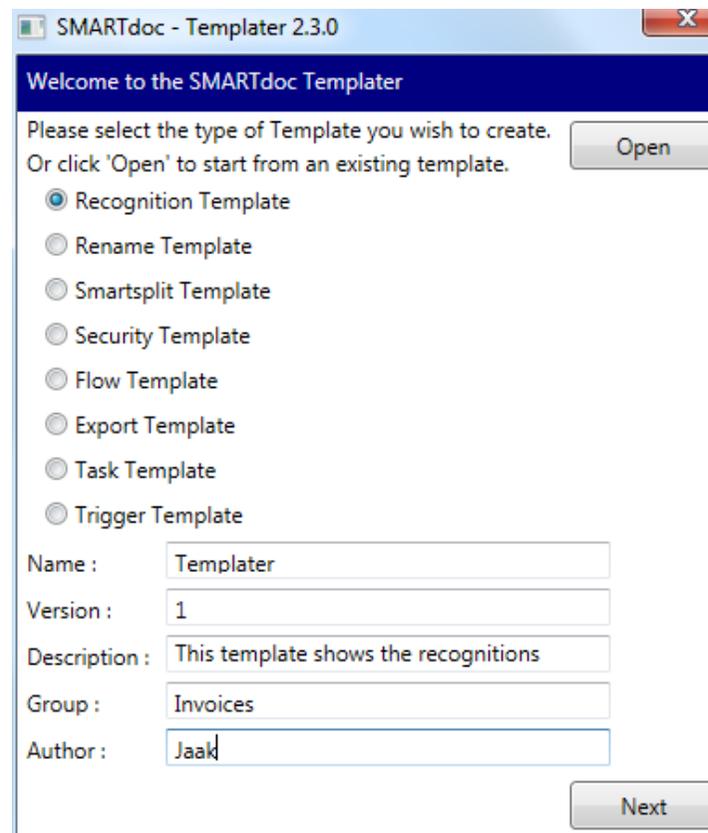
The purpose of a recognition template is to make the document recognizable by SMARTdoc and if desired to fetch information out of the document in indexes.

The more recognition you can build in a template, the better the result will be. The quality of the document to be recognized is also very important: bad quality scans give bad results. Be sure to test your template before you start using it.

CREATE A NEW RECOGNITION TEMPLATE

In the Templater window, choose 'Recognition Template'.

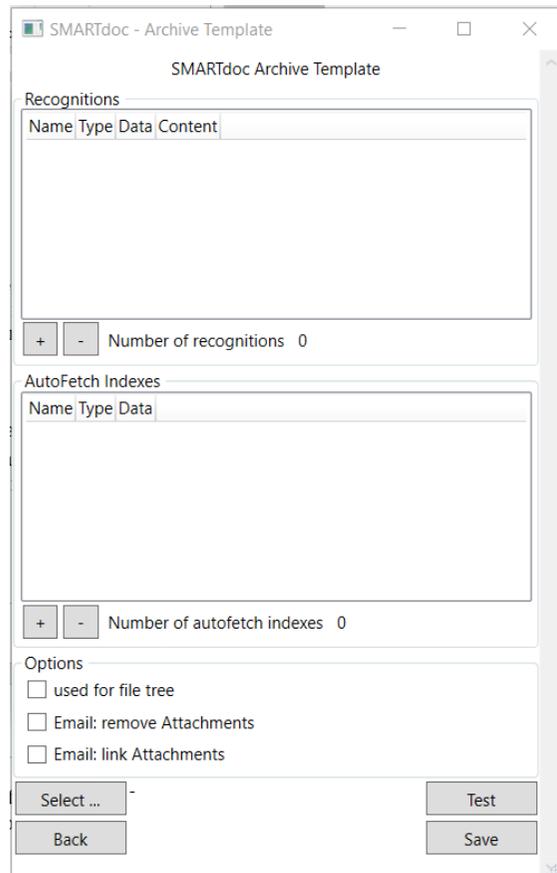
Fill in the name, version, description, group and author of the template. The group is useful because it is used in the search window also.



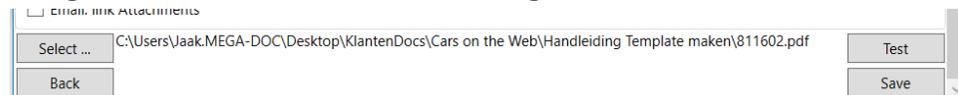
Click 'Next'.

SMARTdoc Templater: Recognition Templates

PARTS OF THE RECOGNITION TEMPLATE



At the bottom of the screen you see a 'Select' button. This is an important button, because here you select an example of the document to be recognized, it will be used to test the recognitions and the autofetches.



The upper part of the screen is used to define how the document will be recognized; the second part is used for the indexes you want to fetch. First we select an example, and then we click on '+' in the Recognitions.

Important remark: check if the text is readable before you start with the template. If the document is an image send it through an OCR program to make the text readable.

RECOGNITIONS

Recognition always has two parts: the value to look for and the expected result. Every recognition has a unique name. You can't have two recognitions with the same name.

We will walk through the different possibilities of recognition.

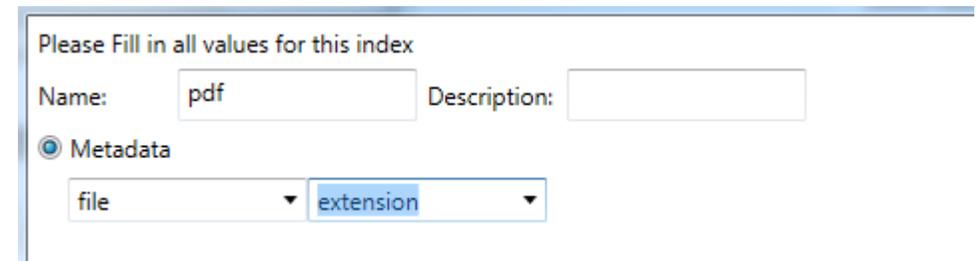
METADATA: EXTENSION

The upper part of the recognition contains the metadata linked to the file.

You have metadata connected to the file, image or user.

This means that you can connect recognition with the user logged in at that moment.

We use the example of the extension:



EXPECTED VALUE

Now we have indicated the kind of recognition, we have to define what the extension is we expect: this is done at the bottom part of the templater.



We have different options:

- Contains: means the typed word is part of the found result
- = (text): the text is exactly the same as the found result
- Starts with, ends with: the found text has to start with or end with the expected value typed
- Exclude: the types expected is not part of the found text
- <, <=, >, >=: are all used for numbers

SMARTdoc Templater: Recognition Templates

We can indicate if our value is case sensitive or not.

After defining the result we save the recognition index.

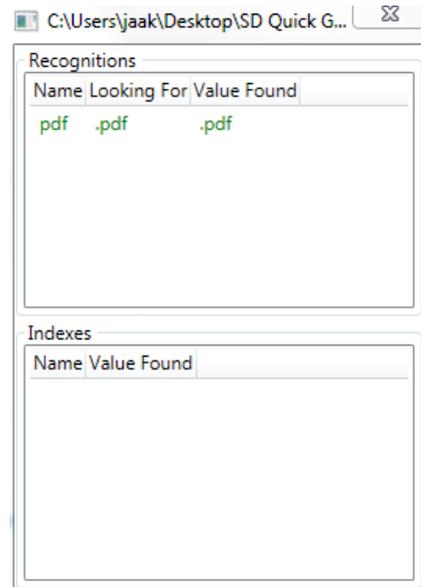
TEST THE RECOGNITION

Each time we define a recognition we test to see if it works.

To do that we click the button 'Test' at the button of our Templater screen.

SMARTdoc asks for a document to test.

If the test is fine, the result is shown in green, otherwise it is shown in red.



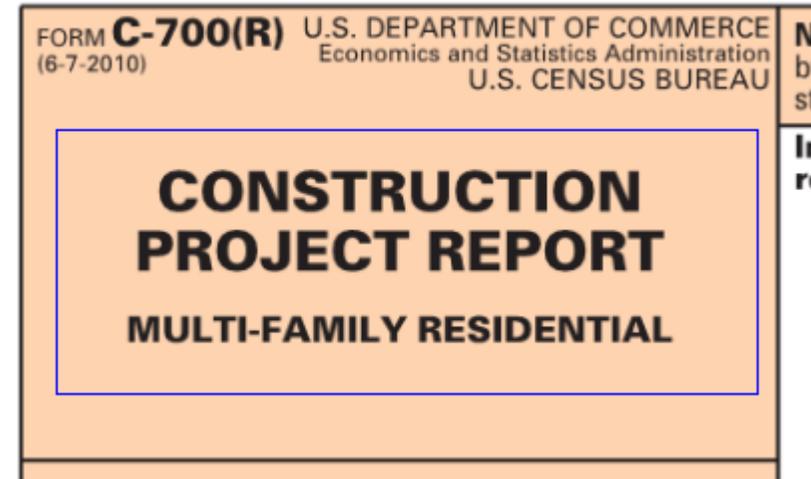
In our example, everything is fine, we can continue with the other recognitions.

TEXT RECOGNITION: LOCATION

When you have a searchable PDF or a Word document, you can fetch text out of it and check if it is recognized.

In the Templater, you create a new recognition and choose: 'Text' 'Location'; be sure that the pdf example is selected. Click 'Get coordinates'.

In the window that opens, select the rectangle where the text is found:



The Adobe coordinates for the selected text are saved:

Location

StartPoint: X: 33 Y: 936

EndPoint: X: 210 Y: 870

At the expected value line, type the text to be found; in our example it could be: 'Construction Project Report'.

Don't forget to save the index and to test it.

Remark for Excel recognition: in Excel the x and y coordinates are the cells; if the text is found in cell B3: you indicate 'x=2; y=3'.

TEXT RECOGNITION: DIRECTION

It is not always possible to define the region where the text has to be found. OCR text recognition can be hard sometimes.

If needed, you can also use 'Direction'.

Direction works in three ways:

- Left: takes the indicated number of characters before the typed text
- Right: takes the indicated number of characters after the typed text
- Between: takes all the text between the first and second text.

SMARTdoc Templater: Recognition Templates

You can use direction in combination with selection; in that case you don't have to define a number of characters. The recognition stops at the end of the selection.

MAIL RECOGNITION

First of all: an email is easily recognized by the extension. A Microsoft Outlook mail has the extension '.msg'.

Next to that you can use the sender, the receiver, the subject as recognition.

PATH RECOGNITION

The purpose here is that SMARTdoc looks at a specific folder and checks the name.

In the field path you indicate the number of the path. Number 0=the highest level (the drive itself). Counting down, you get 1, 2 etc... Users is always level 1 on a PC.

Recognition on path can be very useful because the name of a folder is always fixed.

An example:

'C:\Users\Jaak.MEGA-DOC\Desktop\KlantenDocs\Cars on the Web\Voorbeelden', the folder 'Voorbeelden' is path 6.

In the templater we indicate path 6

Path

Path:

6

The expected value will be 'Voorbeelden'

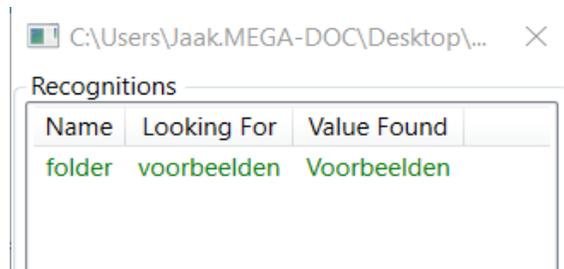
Expected Value:

= (text)

voorbeelden

Is Case Sensitive

When putting a document in this folder and testing the template, we get this result:



BARCODE

SMARTdoc recognizes barcodes and you can indicate that a certain barcode has to be found and can be used to recognize a document.

AUTOFETCH INDEXES

The idea of the autofetch indexes is almost the same as the recognition, but you will not indicate an expected value.

We will pick out the differences between both.

SPECIAL OPTIONS IN THE HEADER OF THE AUTOFETCH

Please Fill in all values for this index

Name:

Description:

Is Workflow Template

Rename file to this index

Not editable

Is Security Template

Mandatory

Autocomplete

Add To AutoComplete Values

Show existing index values

Add to tree view

Hidden

Multi-index

Informative

When creating an autofetch, there are some options available in the header of the index:

- Is Workkflow Template: this means that if there is a workflow with the same name as the fetched value, it will be activated
- Is Security Template: means that if there is a security template with the same name as the fetched value, it will be activated
- Rename file to this index: the file name is changed to the fetched value
- Mandatory: index has to be filled in or the file is not save din SMARTdoc; if the file is saved through an automatic watchfolder the use gets in SMARTdoc a red exclamation mark to show that there is a mssing value
- Not editable: the index value fetched cannot be changed
- Autocomplete: when using an autocompleting list from the database, the rest of the values will automatically be filled in
- Add to AutoComplete values: adds the value to the list used for the autocomplete
- Hidden: an index you don't see, but it is saved in the SMARTdoc database

SMARTdoc Templater: Recognition Templates

- Show existing index values: when you type something in this index, you see the already used values
- Multi-index: you put more than one value in the index
- Add to tree view: when using the old tree view, you see this index
- Informative: the information in this index is not used when searching

METADATA

Works the same as the recognition metadata, with the difference that there is no expected value.

This is also used to fetch the user name as metadata: the name of the person logged in who saved to file in SMARTdoc.

TEXT

Works exactly the same as the recognition part, in autofetch we will often use the combination of location and direction to fetch exactly what we want.

MAIL

When adding mails to SMARTdoc we can fetch the sender, the addressed person, the 'cc', the "bcc" and the attachments

PATH

The path fetches the name of the defined folder. This is used to take over an existing tree structure in SMARTdoc.

FIXED

Is an index where the user can type free text. You can already provide something in there or you can use fixed indexes, that can be connected to this index.

COUNTER

You type the name of the counter, SMARTdoc will add a number each time you archive a document. When documents are deleted the numbers are not reused.

OGM

Used for the structured message in a payment:

+ + + 0 4 6 / 6 3 5 6 / 6 2 0 2 0 + + +

SMARTdoc fetches the numbers between the plusses and can replace the '/' with other characters.

BARCODE

Fetches the value of a barcode found in the document. You can indicate the length of the used barcode to make sure you fetch the right value.

UBL

When using SMARTdoc FlowFin to create UBL files, you define which information from the UBL file you want to use in SMARTdoc

DATABASE QUERY

In here you can type a SQL statement to fetch information directly from the SMARTdoc database

REMOVE WHITESPACE

If there are blanks in the found result, SMARTdoc will remove them. This is good when the original document is not so good and the OCR added extra spaces to text.

IS CASE SENSITIVE

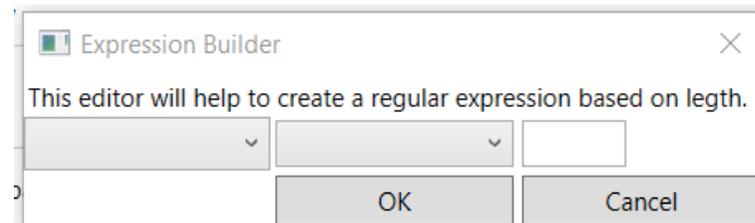
Is only important for recognitions not for fetches. A text is fetched like it is found in the text, with or without cases.

REGULAR EXPRESSIONS

You can indicate the format of the expected value.

For example, you can indicate that the fetched number has to be 8 characters. This is done by a regular expression.

By clicking the Help button you can get some assistance:



By typing the expression yourself, you can go much further:

`\d{4}-(*?)-.*`

SMARTdoc Templater: Recognition Templates

When you need help on building your regular expressions, we are willing to help you.

REPLACE

You can replace spaces or dots in a fetched index by another character or leave it blank:



'\' replaces the dots

'\s' replaces the spaces

'\' replaces the dash

SAVING THE TEMPLATE

Once you are finished, you can save the template.

The template is an XML file. This xml file shows the recognitions and the indexes fetched.

```
<?xml version="1.0" encoding="utf-8"?>
<Template Name="Proximus Invoice" Description="Demo 2013" Type="Archi
  <Indexes>
    <Index Name="Type" Description="" IsFlowTemplate="False" IsSecuri
      <Validation Type="Fixed" IsCaseSensitive="false" Data="Incoming
        <Barcode ExtractBarcode="true" />
      </Validation>
    </Index>
    <Index Name="VAT/TVA/BTW NR" Description="" IsFlowTemplate="False" Is
      <Validation Type="Fixed" IsCaseSensitive="false" Data="02022395
        <Barcode ExtractBarcode="true" />
      </Validation>
    </Index>
    <Index Name="Supplier" Description="" IsFlowTemplate="False" IsSe
      <Validation Type="Fixed" IsCaseSensitive="false" Data="Proximus
        <Barcode ExtractBarcode="true" />
      </Validation>
    </Index>
```

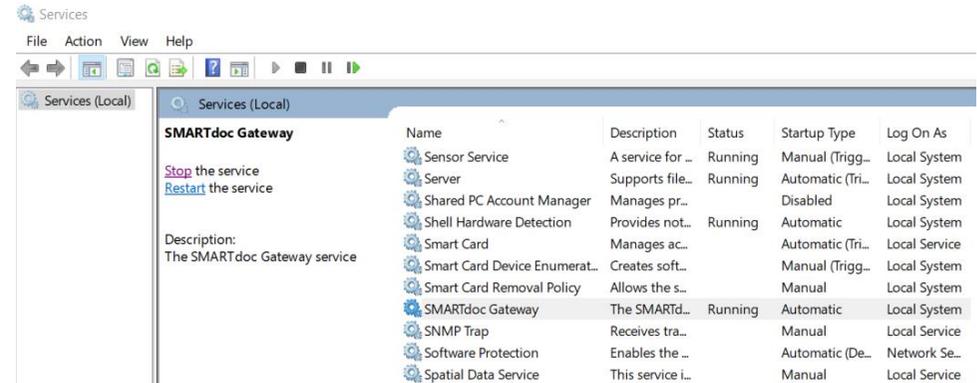
ADD THE TEMPLATE TO SMARTDOC

A template is added to SMARTdoc the same way as a document, drag and drop the template on the SMARTdoc icon. When adding the recognition template, you will be asked for the security and workflow to be used.

RESTARTING THE GATEWAY

When you add templates to SMARTdoc, you have to restart the gateway to make them functional.

The gateway is found on the server in 'Services':

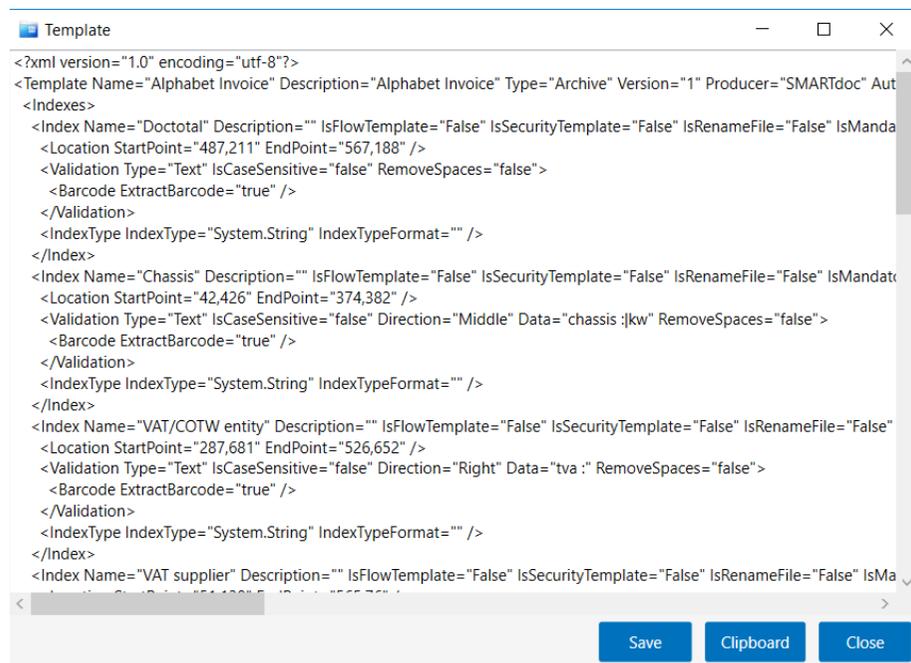


CHANGING THE ORDER OF INDEXES

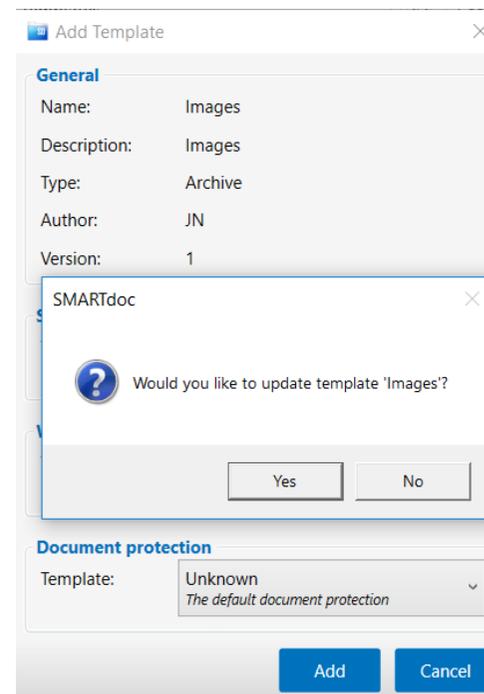
You cannot change the order of indexes in the template, but it is possible in de xml.

Cut and paste the index to the desired location:

SMARTdoc Templater: Recognition Templates



```
<?xml version="1.0" encoding="utf-8"?>
<Template Name="Alphabet Invoice" Description="Alphabet Invoice" Type="Archive" Version="1" Producer="SMARTdoc" Aut
<Indexes>
  <Index Name="Doctotal" Description="" IsFlowTemplate="False" IsSecurityTemplate="False" IsRenameFile="False" IsMandat
  <Location StartPoint="487,211" EndPoint="567,188" />
  <Validation Type="Text" IsCaseSensitive="false" RemoveSpaces="false">
    <Barcode ExtractBarcode="true" />
  </Validation>
  <IndexType IndexType="System.String" IndexTypeFormat="" />
</Index>
  <Index Name="Chassis" Description="" IsFlowTemplate="False" IsSecurityTemplate="False" IsRenameFile="False" IsMandat
  <Location StartPoint="42,426" EndPoint="374,382" />
  <Validation Type="Text" IsCaseSensitive="false" Direction="Middle" Data="chassis :|kw" RemoveSpaces="false">
    <Barcode ExtractBarcode="true" />
  </Validation>
  <IndexType IndexType="System.String" IndexTypeFormat="" />
</Index>
  <Index Name="VAT/COTW entity" Description="" IsFlowTemplate="False" IsSecurityTemplate="False" IsRenameFile="False"
  <Location StartPoint="287,681" EndPoint="526,652" />
  <Validation Type="Text" IsCaseSensitive="false" Direction="Right" Data="tva :|" RemoveSpaces="false">
    <Barcode ExtractBarcode="true" />
  </Validation>
  <IndexType IndexType="System.String" IndexTypeFormat="" />
</Index>
  <Index Name="VAT supplier" Description="" IsFlowTemplate="False" IsSecurityTemplate="False" IsRenameFile="False" IsMa
```



Be sure to select the entire part of an index, starting with `<Index name=" " >` till `</Index>` .

Save the template and put it back in SMARTdoc

REPLACING AN EXISTING TEMPLATE

When you change a template and want to put the new version in SMARTdoc, that has to be done on the SMARTdoc server.

A client gets the remark that the template already exists: on a server you get this message:

After updating the template, you have to restart the template.